

Introduction

In 2011, 2013 and 2016 the National Digital Stewardship Alliance (NDSA) conducted surveys of U.S. organizations currently or prospectively engaged in web archiving to better understand the landscape: similarities and differences in programmatic approaches, types of content being archived, tools and services being used, access modes being provided, and emerging best practices and challenges.

The resulting reports are available here:

2011 NDSA web archiving survey

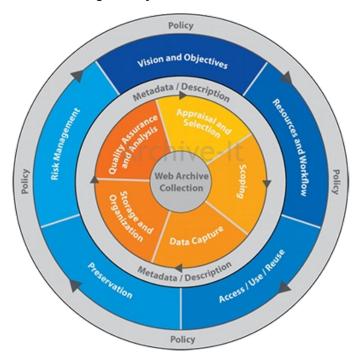
2013 NDSA web archiving survey

2016 NDSA web archiving survey

The NDSA is releasing this updated survey to continue to track the evolution of web archiving programs in the United States. The aggregate responses will be reported to NDSA members and summary results will be shared publicly.

Some questions refer to the Web Archiving Life Cycle model, Here's a reminder of that model's explanation before we get into the survey https://archive-it.org/static/files/archiveit life cycle model.pdf

Web Archiving Life Cycle model

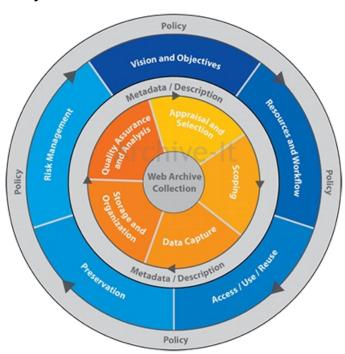




About Your Organization

* 1 . N	Name of your organization
* 2. 7	Type of organization
	Archive
	Historical Society
	College or University
	Museum
	Public Library
	Commercial
	Consortium
	K12 School
	Government: Federal
	Government: State
	Government: Local
	Other (please specify)
3. [Do you or your organization belong to any of these three groups? Choose all that apply.
	Digital Library Federation (DLF) https://www.diglib.org/
	International Internet Preservation Consortium (IIPC) http://netpreserve.org
	National Digital Stewardship Alliance (NDSA) http://digitalpreservation.gov/ndsa
	Society of American Archivists Web Archiving Section https://www2.archivists.org/groups/web-archiving-section https://www2.archivists.org/groups/web-archiving-section

As a quick reminder for the questions below, here is a graphic representation of the Web Archiving Lifecycle model:



* 7. On what dimensions of the Web Archiving Lifecycle Model https://archive-it.org/static/files/archiveit life cycle model.pdf) has your organization made the most progress? Choose three.

	Vision	and	Objectives
--	--------	-----	------------

- Policy
- Resources and Workflow
- Risk Management
- Appraisal and Selection
- Scoping
- Data Capture
- Quality Assurance and Analysis
- Storage and Organization
- Preservation
- Metadata / Description
- Access / Use / Reuse

* 8. On what dimensions of the Web Archiving Lifecycle Model https://archive-it.org/static/files/archiveit life-cycle-model.pdf) has your organization made the least progress?
Choose three.
Vision and Objectives
Policy
Resources and Workflow
Risk Management
Appraisal and Selection
Scoping
Data Capture
Quality Assurance and Analysis
Storage and Organization
Preservation
Metadata / Description
Access / Use / Reuse
* 9. What are the goals of your web archiving activity? Choose all that apply.
Archive your own or affiliated web content (e.g., university archives archiving the university website or state library archiving state agency websites)
Archive content from other organizations or individuals (e.g., research library archiving third-party web content as par of topical collection building or special collections department archiving web content associated with a manuscript donor)
Other (please specify)
10. Optional: briefly share any driving factors which determine your collecting goals:
11. What year did your organization begin archiving web content?
, , , , , , , , , , , , , , , , , , ,

* 12. How much full-time equ archiving? Choose one.	ivalent (FTE) staff time does your organization dedicate to web
	zation's web archiving activity relied upon one dedicated, full-time me of three other full-time employees, you would indicate "more than 1,
.25	
.5	
.75	
<u> </u>	
more than 1, less than 3	
3 or more	
three.	erations for the development of your web archiving program? Choose
	cher interactions, web analytics, use cases)
Cost (e.g., budgeting, service	e allowance utilization, staffing level requirements)
Data volume (e.g., data volu	me collected, objects collected, acquisitions statistics)
Institutional buy-in (e.g., prog	grammatic growth, stakeholder testimonials, resource commitments)
Loss (e.g., link and/or referen	nce rot of archived resources)
Quality (e.g., accuracy, comp	pleteness, replay fidelity)
Risk management (e.g., perr	mission responses, takedown requests, policy conformance)
Other (please specify)	

	What are the top staff skills that are essential to the development and success of web archiving your organization? Choose three.
	Appraisal and selection (e.g., determining what web content to collect)
	Archiving tools (e.g., configuring or operating web archiving tools)
	Collaboration and communication (e.g., advocacy, coordination, marketing, or outreach)
	Domain expertise (e.g., knowledge of subjects that are the focus of web archiving)
	Metadata (e.g., familiarity with metadata standards, cataloging experience)
	Quality assurance (e.g., analyzing and troubleshooting web archive quality issues)
	Software development (e.g., able to develop software or web applications)
	Web technologies (e.g., familiarity with web architecture, design, formats, or platforms)
	Other (please specify)
app	What types of content do you have concerns about your capacity to archive? Choose all that oly. Audio Blogs Databases Interactive media Social media Video Other (please specify)
* 16.	Do you currently collaborate with any other institution on any area of web archiving? Yes
	No
	No, but interested
	Don't know

	If you answered yes to question 16, on which areas of web archiving do you collaborate with other institution? Choose all that apply.
	Best practices for policy and risk management
	Capture configuration and optimization
	Collaborative collection development
	Input on APIs and standards
	Metadata standards and application
	Quality assurance techniques and strategies
	Tool development, documentation, or user feedback
	Other (please specify)
* 18.	In what areas of web archiving are you most interested in collaborating? Choose all that apply.
	Best practices for policy and risk management
	Capture configuration and optimization
	Collaborative collection development
	Input on APIs and standards
	Metadata standards and application
	Quality assurance techniques and strategies
	Tool development, documentation, and user feedback
	Other (please specify)
* 19	What barriers do you face for collaborating on web archiving? Choose all that apply.
	Still in planning / pilot stage; not much to share
	Lack of institutional support
	Lack of time to spend on collaborating
	Using a proprietary system
	Institutional policies prohibit or limit collaboration
	Other (please specify)



Tools and Service Providers

20.	If capturing locally, what tool(s) do you use? Choose all that apply.
	Adobe Web Capture
	Grab-a-Site
	Heritrix
	HTTrack
	Teleport Pro
	Web Archiving Integration Layer (WAIL)
	Web Curator Tool
	WebRecorder
	Wget
	Other (please specify)
21.	Are you capturing social media utilizing API tools?
	Yes
	No
22.	If you answered yes to question 21, please list the tools you use

	If you are (or have been) using an external service for data capture, which one(s) do you use? loose all that apply.
	Archive-It
	Hanzo Archives
	Internet Archive's contract crawling services
	OCLC Web Harvester
	Other (please specify)
	If you are (or have been) using an external service for data capture, have you replicated any of r data to another repository?
	Yes
	No
	If you answered yes to question 24, to what kind of repository have you replicated the data?
	Local repository
	External preservation service provider
26. app	If you have not replicated any of your data to another repository, why not? Choose all that ly.
	Trust web archiving service provider
	Building local infrastructure
	No place to store / maintain it
	Not sure what we'd do with it once we got it
	Other (please specify)



Access and Discovery

2	. What kind(s) of access does your organization itself provide? Choose all that apply.
	URL search
	Full-text search
	Browse list by URL
	Browse list by title
	Catalog records: collection-level description
	Catalog records: item-level description
	Finding aids
	Application programming interfaces (APIs)
	WARC files
	Derivative datasets
	Other (please specify)
* 28	3. Do you have active researchers utilizing your web archives?
	Yes
	No No
	I don't know
	O. If you answered yes to question 28, could you provide a summary of how researchers are using our web archives?

-	• • • • • • • • • • • • • • • • • • • •	ding notifying and seekin o their web content. (cho	g permission from content ose one for each row)
	No action	Notifying	Requesting permission
Capturing			
Providing restricted access	\bigcirc	\bigcirc	\bigcirc
Providing public access		\bigcirc	
Other (please specify)			
31. Have you ever had a made accessible without Yes No If yes, please specify	•	ing or to take down conte	nt that you've crawled or
	ND:	SA%	
	2017 NDSA Web	Archiving Survey	
chiving Policies			
		ding notifying and seekin o their web content. Choo Notifying	g permission from content ose one for each row. Requesting permission
Capturing			
Providing restricted access		\bigcirc	
Providing public access			
Other (please specify)			

38. If you selected "Sometimes <i>l</i> it depends" for question 37, under what circumstances does your organization ignore robots.txt? Choose all that apply.
If your organization owns the copyright/has some other special access right (e.g., state archive collecting state agency websites)
If permission is secured or appropriate notices have been sent
In order to capture essential content (e.g., stylesheets, images, etc.)
Other (please specify)
39. What resources has your organization relied upon in the development of its own copyright and access policies? Choose all that apply.
ARL Code of Best Practices in Fair Use for Academic and Research Libraries http://www.arl.org/focus-areas/copyright-ip/fair-use/code-of-best-practices
Oakland Archive Policy http://web.archive.org/web/20140812200246/http://www2.sims.berkeley.edu/research/conferences/aps/removal-policy.html
Section 108 Study Group Report http://www.section108.gov/
Consultation with legal counsel
Statutory authority
Web archiving policies and practices of other organizations
Previous NDSA Web Archiving Survey reports
Other (please specify)

If you are willing to share your web archiving policies, please send a copy or provide a link to ndsa@diglib.org.



2017 NDSA Web Archiving Survey

Thank you for participating in the 2017 NDSA Web Archiving Survey!

Results will be shared in 2018.